

ABOUT

I am an M.S. student in Computer Science at **Institut Polytechnique de Paris (IP Paris)**, with a B.S. in Data Science and Big Data Technology from **Southern University of Science and Technology (SUSTech)**. I am currently a Research Assistant at **SUSTech-NLP**, advised by Prof. Guanhua Chen.

My research interests broadly lie in **efficient LLM systems** and **agentic AI**. I have worked on **LLM-SLM model collaboration** for efficient routing and inference, and I am currently working on **efficient agentic exploration** for test-time adaptation. I am also interested in **LLM for health**, as an application domain where low-resource constraints and trustworthiness are both critical.

EDUCATION

Institut Polytechnique de Paris

M.S. in Computer Science

Paris, France

Sep 2025 – Jul 2027

Southern University of Science and Technology (SUSTech)

B.S. in Data Science and Big Data Technology

Shenzhen, China

Sep 2021 – Jul 2025

RESEARCH

SUSTech-NLP Group

Research Assistant, advised by Prof. Guanhua Chen.

Shenzhen, China

Jan 2025 – Present

- **LLM-SLM Collaboration:** Proposed RouterXBench to decouple router ability from scenario alignment; systematically compared routing signals and showed that internal hidden states capture model uncertainty more effectively than output logits or external embeddings, leading to ProbeDirichlet with learnable Dirichlet aggregation across layers.

INTERNSHIP

Institute of Intelligent Medical Research (IIMR), BGI

Machine Learning Intern (AI for Genomics).

Shenzhen, China

Jul 2024 – Jan 2025

- Developed cfDNA-based feature representations via statistical filtering and model-guided selection; applied ensemble learning for early colorectal cancer detection, achieving **94.88% sensitivity**.

Peking University Shenzhen Hospital

Data Analyst Intern.

Shenzhen, China

Jul 2023 – Oct 2023

- Performed biostatistical analysis on multi-timepoint clinical records to compare postoperative outcomes across surgical methods.

PUBLICATIONS

Towards Fair and Comprehensive Evaluation of Routers in Collaborative LLM Systems



- Wanxing Wu*, He Zhu*, Yixia Li*, Yun Chen, Guanhua Chen[†] (*Equal contribution)
- Under Review, 2026 [[arXiv:2602.11877](#)]

ACADEMIC PROJECTS

ReflexSQL: A Self-Reflective Agent for Robust Text-to-SQL Generation

Dec 2025 – Jan 2026

QueryPilots, Institut Polytechnique de Paris [[Poster](#)]

- Built an execution-aware self-reflection framework for enterprise-scale Text-to-SQL on Spider 2.0-Lite, combining candidate voting, execution feedback, and memory-augmented exploration.

Conservative Off-Policy Trust Region Policy Optimization

Nov 2025 – Jan 2026

- Developed a dimension-wise masking strategy for off-policy TRPO that leverages on-policy gradient alignment to suppress conflicting update components, ensuring update directions remain consistent with the on-policy objective; validated on PyBullet/Gym locomotion tasks.

HONORS & AWARDS

- **Outstanding Graduate**, Zhiren College, SUSTech (Jul 2025).
- **Outstanding Graduation Thesis**, Top 5%, SUSTech (Jul 2025).
- Outstanding Student Scholarship, SUSTech (Oct 2024).
- **MCM/ICM Mathematical Contest in Modeling**, Honorable Mention (Feb 2024).

SKILLS

LLM & ML: inference pipeline (vLLM, SGLang), agentic-RL workflows, SFT/RL training, model evaluation.

Programming & Dev: Python, PyTorch, Docker, Linux, Git, shell scripting.

Languages: English (IELTS 6.5, CET-4, CET-6); Chinese (Native).